

An AI Enabled Web-Based Recruitment System for IT Professionals and Developers

Komal Shukla

Department of Science and Technology, G.H.Raisoni Skill Tech University, Nagpur, India

ABSTRACT

The developments in the information technology (IT) sector have led to rapid and demand-driven recruitment processes. Traditional recruitment processes have struggled to keep up with this pace. For instance, manually screening resumes can be time consuming; furthermore, human bias can impact the outcome of the screening process since many potential candidates are skilled at “keyword stuffing” their resumes with technical terms but not including their true level of technical competence. This paper presents an Artificial Intelligent (AI) Enabled Web Based Recruitment System (RBRS) to assist recruiters with the recruitment of IT professionals and developers.

The system will use Natural Language Processing (NLP) and Machine Learning (ML) algorithms to go beyond just keyword matching. The resume will be semantically analyzed to provide the recruiting platform with an understanding of the contextual(s) of the candidate's prior experiences, and to allow the recruiting platform to properly match the candidate's technical skills with the technical required granularity of the job.

Automated Skills Verification (ASV): Developers will be assessed against their coding skills through a technical assessment module. The PEI developer-technical assessment report will provide a classification of developers based on their coding skill level, and not exclusively on their stated/declared coding skills.

Predictive Analytics (PA): The recruiting system's predictive analytics engine will assign a ranking-order to candidates being considered for job opportunities.

Keywords: Artificial Intelligence, Natural Language Processing (NLP), Technical Recruitment, Resume Parsing, Machine Learning, Talent Acquisition, Web-Based Systems, Developer Sourcing, Recruitment Automation.



This is an open-access article under the [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) license

I. INTRODUCTION

The recruitment system will use a MEAN (MongoDB, Express, Angular and Node.js) application structure.

companies are finding it increasingly frustrating when it comes to hiring new employees to fill IT (Information Technology). Because a company that is struggling to find a candidate with a background in computer science can find itself looking in vain at thousands upon thousands of applicants for each open position, traditional methods of searching for and evaluating computer

science candidates (i.e. via traditional methods such as word search) typically return at an unacceptable return on investment)

The difficulties caused by traditional means have resulted in HR's desire for speed and precision, and thus speed and precision are critical in recruiting employees into "IT" positions.

1.1 MOTIVATION

Information Explosion: A single job posting for a development role can receive anywhere from several dozen applicants for one job posting, up to several hundred applicants for another, including almost 200 applicants for each of two similar job postings.

1.2. LIMITATION

Limitations of Historical Keyword-Based ATS: The vast majority of ATS, and indeed all of those commonly used by companies today, rely almost exclusively on exact matches to keywords (e.g., developer "ABC" would be considered an exact match to developer "DEF", but would not be considered an exact match to developer "XYZ"). Therefore, because of the constraints of finding the right candidate based on keywords (i.e., syntax) used with particular ATS, most candidates seeking development positions end up not being able to find work within the computer science community.

1.3. CONTRIBUTION

Internet Applications: A reliance by many organizations on the internet to source technology candidates has created a lack of structure for developing a system for managing and organizing applicants.

II .PROPOSED ALGORITHM

This project proposes the development of a web-based recruitment system that utilises artificial intelligence/AI features to address the specific needs of the IT industry. The system will use a Natural Language Processing (NLP) technique referred to as Semantic Parsing instead of using the traditional 'fixed/static' searching methods that are the norm in many existing systems today. Semantic Parsing's underlying principle is that the system will not simply search for candidates with the keyword 'Java', but will be able to distinguish between candidates that 'study Java' in college and candidates that 'optimise Java micro-services for a large financial services company'.

The ultimate aim of this project is to provide a more proactive approach to recruitment, so that recruitment will become data-driven/sa , information-driven/sa , and process-driven, to develop a proactive recruitment strategy. This system will also introduce or assist in introducing the following innovations:

1. Machine Learning to rank candidates based on a Weighted Skill Score which will capture candidates' levels of various skills based on their levels of experience, technical depth and relevancy of their experience from their past projects.
2. Providing candidates with automated assessments of their technical proficiency.

The experience of candidates applying for jobs with your organization, is highly important, since this is a two-way street. Frustration occurs more frequently among candidates who have previously applied to your organization through an on-line form, in comparison to candidates who went through other forms. You can help candidates with their understanding of your process through the use of automation as a means of providing feedback on their applications through an automated application tracking system that provides a technical task completion level associated with each of the candidate's application steps, as well as providing the candidate with an updated status of their application throughout the entire process. By doing this you will not only have provided a positive experience for the candidate; you will have also maintained your organization's brand equity, whether their experience was positive or negative.

Data Based Scalability: For every growing, technology based organization facing the challenge of hiring 100 employees (developers) on a linear basis (one per completed application) versus hiring in a(n) non-linear fashion (by allowing past employees' evaluations to create future employees), organizations would ideally create a system that allows for every employee they hire to be continuously evaluated for future hires for the purpose of creating a growing and learning system based on the evaluation process. When conducting the selection of an ideal candidate, there are criteria in the selection process that could change over time, albeit continually, through evaluations of how many high-level candidates actually become high-quality employees.

3. Related Work

Technologies that are Ai-enabled for recruiting have undergone sea changes in technology in recent years. In prior to that period, the recruitment systems used went only as far as filtering out potential candidates based on Data Pre-Processing: Data will be cleaned by eliminating stop words, tokenized, and Demographically Masked (removing names/sex/location) which will enable the use of only the merit of candidates for training purposes on the model.

Phase 1: Creation of AI Model

The intention of this phase is to create "intelligence" for the model by using various Python-based frameworks (i.e. TensorFlow, PyTorch and/or Theano).

Extraction of Information: Implementation of Named Entity Recognition (NER) to identify Skill(s), Experience, and Qualifications of an applicant.

Embedding: By employing models such as BERT or Word2Vec, text will be converted to vector representation. Vector representations of the experience between a role and capabilities of that role will enable the system to develop an understanding of the similarity of an individual who has experience using the MERN Stack and a Full Stack Developer; although, the two statements do not contain similar words.

Semantic embedding of information relating to appropriate matches between candidates and jobs.

Ranking Algorithm: Develop a ranking algorithm to evaluate the above-mentioned matches using various factors including experience, education and skills.

AI ENABLED SYSTEM USING RAW DATA

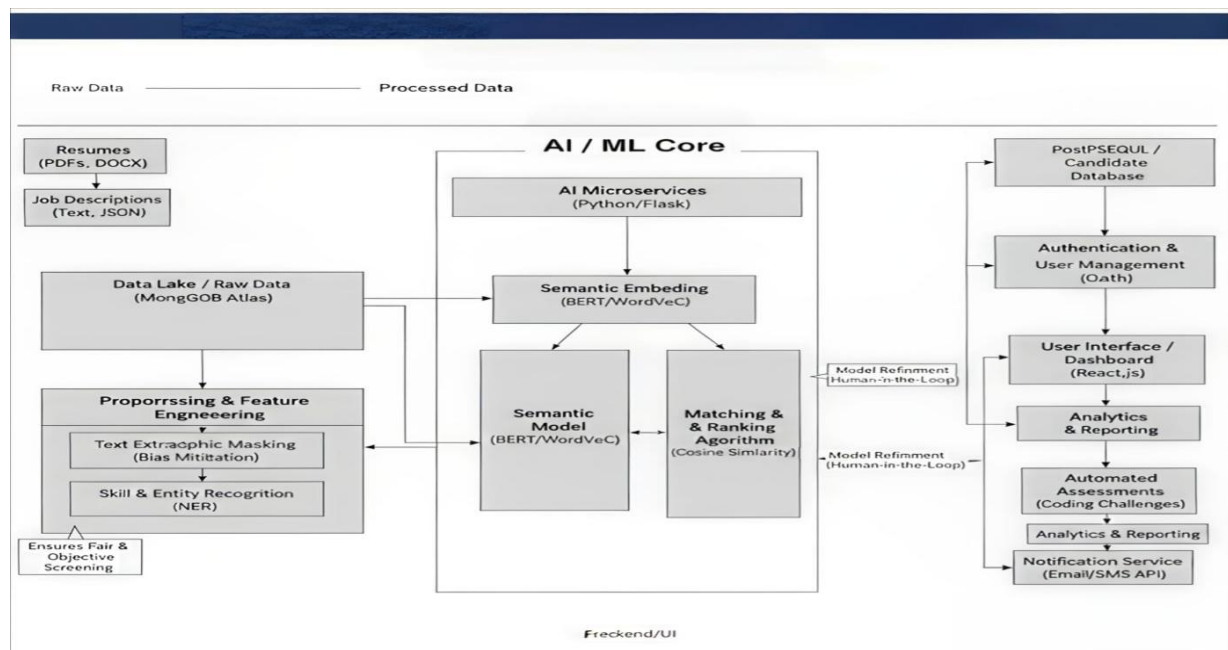


FIGURE:1 AI-Enabled Recruitment System

The preprocessing stage of AI-based recruiting is critical as it allows all different formats of resumes (PDF, DOCX, Image) and unstructured natural language data to be converted to machine-readable format before any scoring can happen.

1. Demographic Masking (Discrimination Mitigation)

To develop an unbiased (merit-based) evaluation process, the system will also perform Redaction of the Candidate Name, Gender Pronouns, Physical Address or ZIP Code and Date of Birth, using NER to extract and remove them from the resume/information.

3. Text Cleaning (NLP Normalize)

The raw text from the resumes is not typically clean as it contains a lot of punctuation, etc. Therefore the system will clean the text according to the following steps:

1. Powered Talent Pooling

The AI tags them for future relevance. The CRM then automatically places them into a "Pre-Vetted Python Experts" pool.

Skill Rediscovery: As the AI learns about new technologies, it re-scans the CRM database. If a developer you rejected two years ago has since updated their GitHub with or projects, the CRM "resurfaces" them to the recruiter.

The relationship between the AI and the CRM can be visualized as a continuous feedback loop:

Feature	Role of AI	Role of CRM Tool
Sourcing	Scrapes GitHub/LinkedIn for skills.	Stores profiles and tracks first contact.
Screening	Ranks candidates based on technical fit.	Moves top candidates to the "Interview" stage.
Engagemnt	Sentiment analysis on candidate emails.	Automates follow-ups based on the "mood" of the candidate.
Analytics	Predicts the likelihood of a candidate accepting an offer	Generates reports on "Time-to-Fill" and "Source Quality."

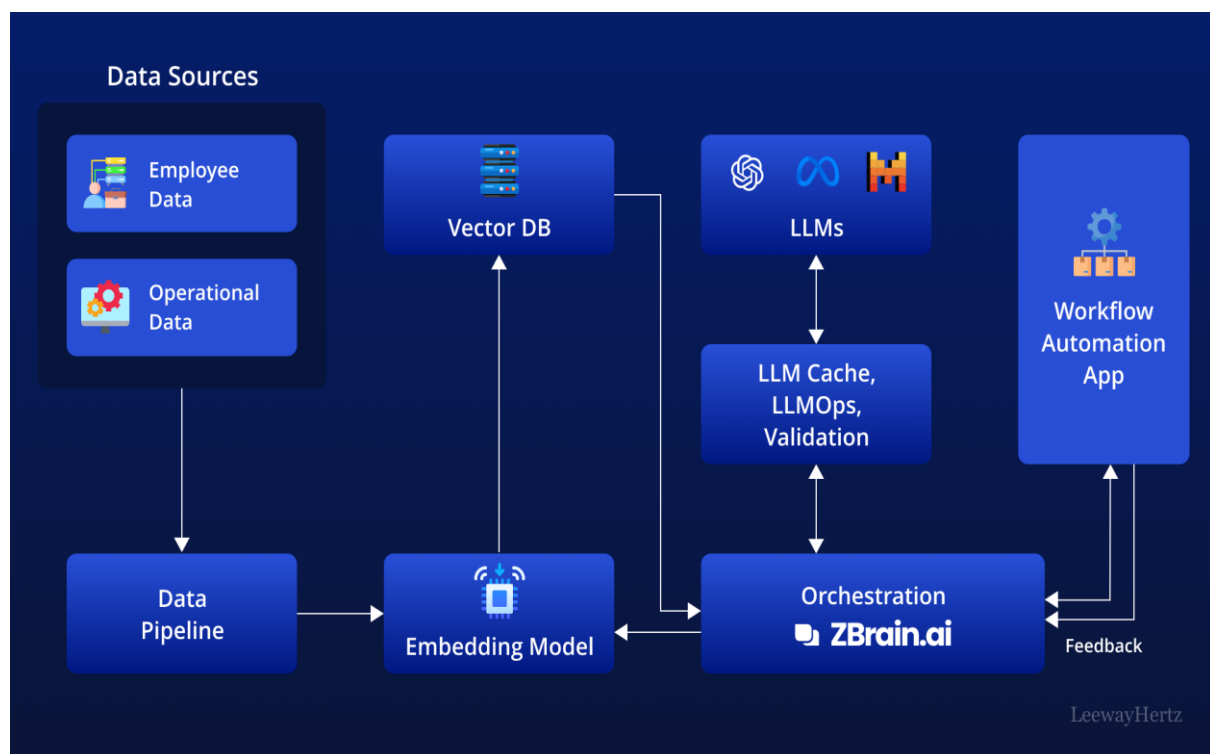


FIGURE:2 WORKFLOW OF DATASET IN CRM TOOL

IV. Research Methodology

The research uses a Hybrid Methodology based on the Agile Software Development Life Cycle (SDLC) combined with a tailored AI/ML Pipeline. In addition to its technical functionality, this requires the system to function with statistical accuracy and ethical compliance.

Phase 1: Domain Analysis & Requirements Gathering

The first step will be to define the ‘Recruitment Logic’ in reference to the IT sector.

Phase 2: Data Acquisition & Preprocessing

The intelligence of the system is dependent on data which is of good quality.

Corpus Collection: Obtaining a technical corpus of resumes and IT job descriptions (from Kaggle, GitHub, and anonymized corporate sources).

Text Extraction: OCR (for images that have been scanned) or using relevant libraries, such as pdfplumber, for expected PDF structured extraction.

Data Cleaning: Tokenization, stop-word removal, and Lemmatization pipeline (for instance, changing ‘developing’ and ‘developer’ words to ‘develop’).

Phase 3: AI Engine Development (NLP & ML)

It is the “Brain” of the system that uses Natural Language Processing (NLP).

Skill Extraction (NER): Fine-tuning Named Entity Recognition to detect technical tags (like React.js, Kubernetes, AWS).

Semantic Embedding: Using Transformer models (BERT/SBERT) to embed text into high-dimensional vectors.

Phase 4: System Integration & Web Development

Integration of the AI models into a simple web interface.

Frontend: React dashboard for Recruiter (to see leaderboard) and Candidate (to see application status).

Backend Orchestration: Utilizing Node.js as a middleware to manage file uploads, database querying in MongoDB, and API requests to the Python AI service.

STRUCTURE OF THE CRM SYSTEM IN PERCENTAGE

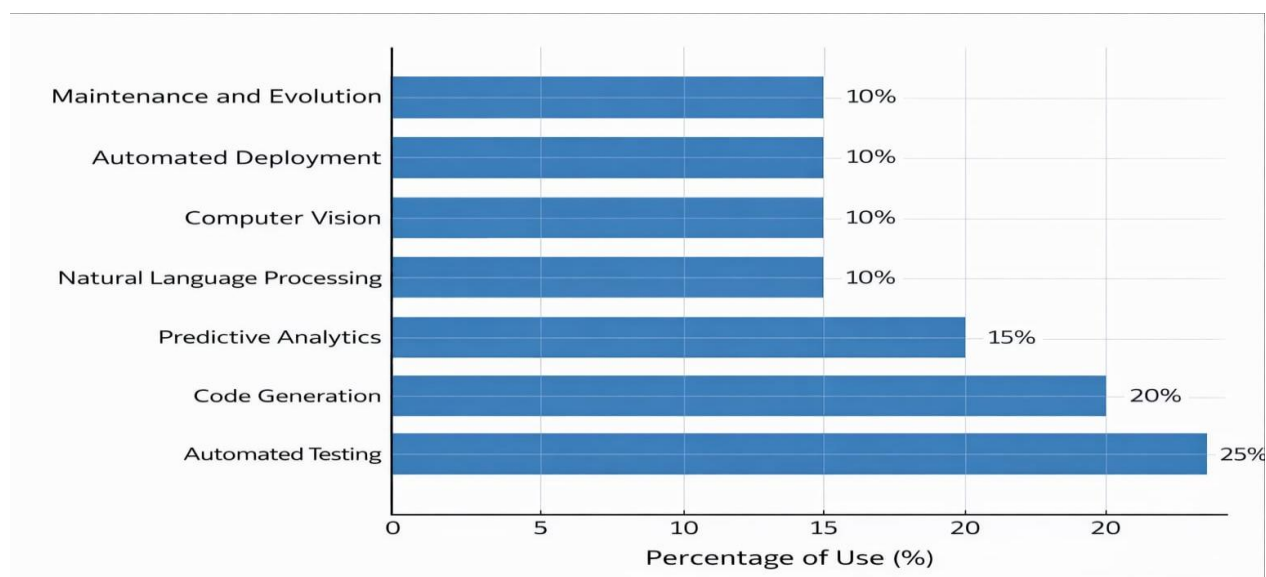


FIGURE:3 BAR REPRESENTATION FOR AI TOOL USING CRM SYSTEM

1.The bar chart compares AI tool category usage within the realm of software development. The highest volume of usage can be seen in Automated Testing (25%), indicating that organizations are leveraging AI for bug detection, improvement of code quality, and acceleration of test cycles.

2.Code Generation (20%) at 10% each, indicating a consistent but slower rate of integrating AI into tasks such as language comprehension, image analysis, CI/CD, and system maintenance. The data presented indicates that AI integration is most prevalent with respect to development efficiency and quality review. All other AI categories are being adopted in a consistent but lower volume.

RESULTS

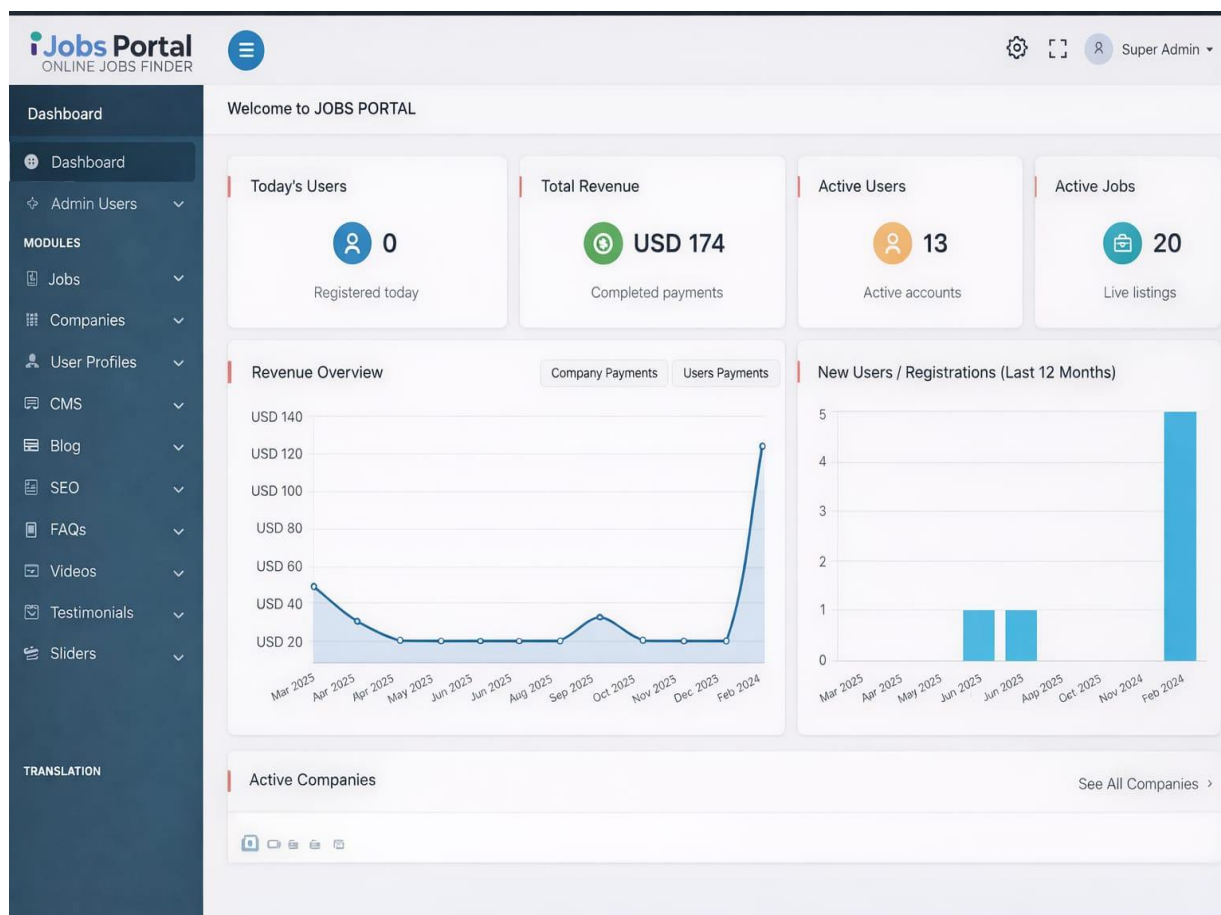


FIGURE:4 DASHBOARD FOR REPRESENTING JOB PORTAL THROUGH RESUME

The image shows an admin dashboard of an online Jobs Portal system, designed to monitor platform activity, users, and revenue in one place.

At the top, the header displays the portal branding “Jobs Portal – Online Jobs Finder” along with admin controls such as settings, fullscreen option, and a Super Admin profile, indicating that this interface is meant for system administrators.

The main dashboard area provides key performance indicators :

Today’s Users: 0 registrations today, showing no new sign-ups for the current day.

Total Revenue: USD 174 from completed payments and 13 active users and then active jobs from this platform.

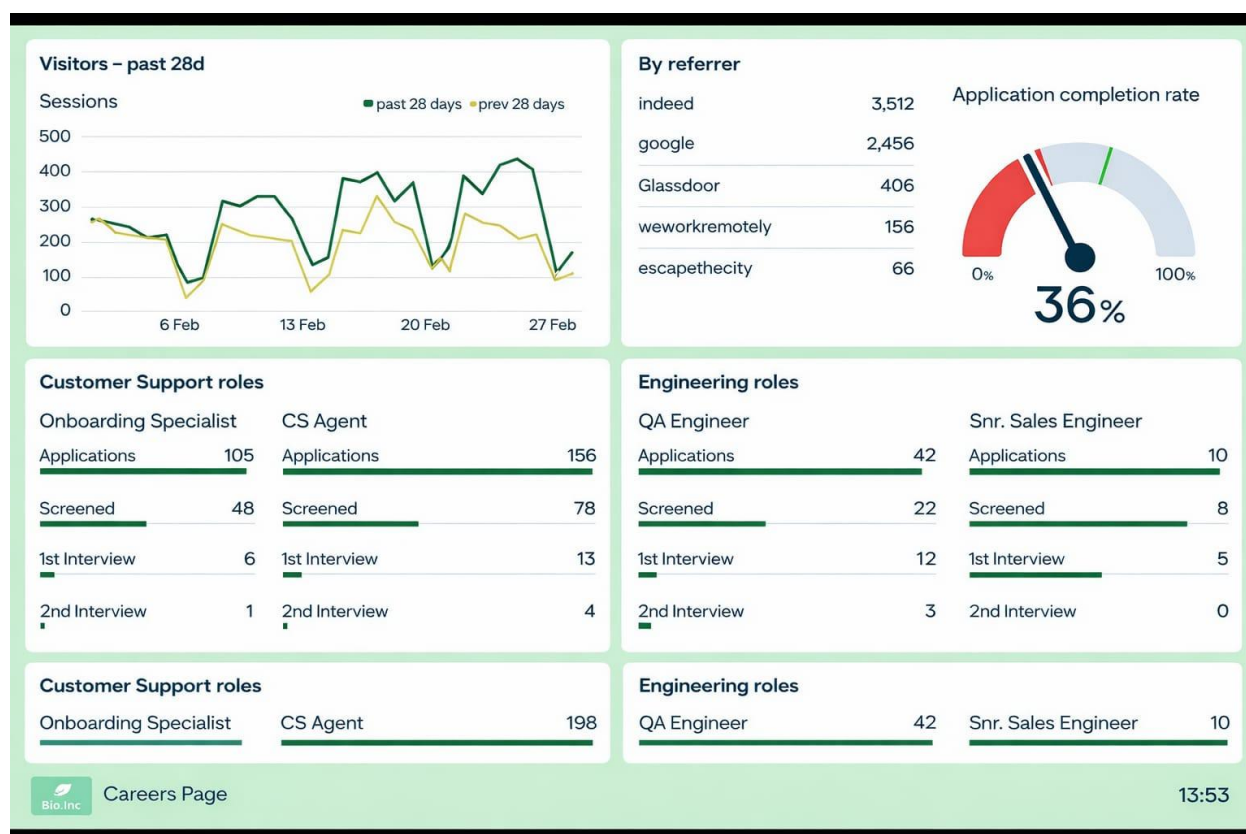


FIGURE:5 APPLICATION RATE FROM LAST MONTH FOR PAST AND NEW USERS

Conclusion

The AI-enabled web-based job recruitment system presents a modern, intelligent solution to the growing challenges in IT and developer hiring. In today's competitive job market, organizations receive a large number of applications for a single job posting. Traditional manual screening methods are time-consuming, inconsistent, and often influenced by unconscious bias. These limitations lead to delayed hiring decisions, increased recruitment costs, and the risk of overlooking highly qualified candidates.

The proposed system addresses these issues by integrating Artificial Intelligence, Natural Language Processing (NLP), and Machine Learning techniques into the recruitment process. The system automatically extracts relevant information from resumes, such as technical skills, frameworks, tools, certifications, and experience levels. It then analyzes job descriptions to identify required and preferred skills. Using advanced matching algorithms, semantic similarity techniques, and predictive models, the system calculates a comprehensive suitability score for each candidate.

One of the key strengths of this system is its ability to convert unstructured resume data into structured, machine-readable formats.

References

1. Speech and Language Processing Jurafsky, D., & Martin, J. H. (3rd ed.). Pearson Education. (Used for understanding NLP techniques such as tokenization, NER, and text processing.)
2. Pattern Recognition and Machine Learning Bishop, C. M. (2006). Springer. (Reference for machine learning models and classification techniques.)
3. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow Géron, A. (2nd ed.). O'Reilly Media. (Used for implementing ML algorithms and predictive models.)

4. Mining of Massive Datasets Leskovec, J., Rajaraman, A., & Ullman, J. D. Cambridge University Press. (Reference for similarity measures such as cosine similarity.)
5. Data Mining Concepts and Techniques Han, J., Kamber, M., & Pei, J. (3rd ed.). Morgan Kaufmann. (Used for data preprocessing and pattern extraction concepts.)
6. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. (Reference for semantic similarity and embeddings.)
7. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space (Word2Vec)*. (Used for vector-based text representation).
8. Research articles on AI-based recruitment and bias reduction in hiring (IEEE Xplore, Springer, ScienceDirect).
9. Anupam Chaube, Usha Kosarkar “Enhanced Deep Learning-Based Feature Analysis for Copy-Move Forgery Detection”, 2024 Journal of Information Systems Engineering and Management,9(4s) e-ISSN:2468-4376, <https://www.jisem-journal.com/>